Cyber Data Websites Related to the talk: Some Statistical Applications in Cyber Security WSS October 17, 2017

David J. Marchette

Below is a set of miscellaneous web sites where one can obtain data related to various topics in cyber security. This is not a comprehensive list, but it does cover a number of different types of data.

It is important to stress that one should always take care in accessing these and similar data. Make sure your security rules allow you to have such data sets on your system, and **never** run any code you obtain from these sites except on a disconnected system under controlled conditions.

For those who want to collect their own network data (with permission of the appropriate network/computer security individuals, of course), check out the site `http://sectools.org/tag/sniffers/`, which contains links to a large number of network sniffers. Note that in some organizations these are forbidden, so make sure you have permission before you download them.

- `http://csr.lanl.gov/data`. A number of (anonymized) data sets collected on the Los Alamos National Laboratory's network.

- `https://www.kaggle.com/c/malware-classification`. The Kaggle malware family data mentioned in the talk.

- `http://www.secrepo.com`. A reposited of a large "somewhat currated" list of security related data.

- `http://networkdata.ics.uci.edu/data.php?id=110`. A social network of hackers at a conference.

- `http://www.usma.edu/crc/sitepages/datasets.aspx`. Data sets available from the US Military Academy (West Point) consisting of various log files from their network.

- `https://stratosphereips.org/category/dataset.html`. A site that captures and disseminates normal data and malware data. There are a lot of files here!

- `http://csmining.org/index.php/malicious-software-datasets-.html`. Windows system calls for normal and malware programs. Also links to some standard data sets such as the Enron email spam data.

- `http://vizsec.org/data`. A list of links to a number of data sets relevant to computer security.

- `https://zeltser.com/malware-sample-sources`. A few links to freely available malware data.